EVALUATING EFFICACY OF THERAPY TECHNIQUES IN CHILDREN WHO STUTTER

Ooi Chia Ai[1], J. Yunus[2]

Department of Electronic Engineering, Universiti Teknologi Malaysia, Skudai, Johor, Malaysia

[1] OoiChiaAi@yahoo.com

Abstract-**This paper presents a computer-based evaluation system for stuttering therapy. The system provides guidance for clinician to determine suitable technique for each of the children who stutter (CWS). CWS respond in unique ways to different therapeutic approaches. The technique that works so dramatically for one of the CWS does not necessarily work dramatically, or at all, for other CWS. Stuttering is so variable and so highly individualized that, few would disagree, no one method works for all CWS. Normally, 2 to 3 months are required to determine suitable techniques for each of the CWS. The effectiveness of each approach depends on the receptiveness of CWS. This means that clinician has to try every technique in order to determine the appropriate techniques. Our software was developed based on standard fluency shaping techniques used in fluency rehabilitation regimen. Digital Signal Processing techniques were implemented to analyze speech signals. Our hypothesis is by doing comparison between CWS's and clinician's average magnitude profiles, the computational analysis help clinician to determine suitable therapy techniques in a faster way.**

Keywords - **Stuttering, children who stutter, therapy techniques, efficacy**

I. INTRODUCTION

Stuttering is the intermittent impairment of fluency and speech rate, pitch, loudness, inflectional patterns, articulation, facial expression and postural adjustments of the client in the absence of word finding problems, speech motor disorders or voice problems. It is known as stammering in Britain. The term stuttering will be used herewith. The earliest age of onset of stuttering is about 18 months (when speech emerges) and the latest age of onset vary from 7-13 years of age. The incidence of stuttering (which is the approximate percentage of the population who have stuttered at any time in their lives) can be as high as about 10% [1].

While it is important for both adults and children to have speech disruptions assessed for evidence of stuttering, this is particularly crucial for children in the critical speech development years of 2 to 4. The consensus now is that stuttering should be treated as early as possible, primarily because it becomes less tractable as children get older. This is presumably because neural plasticity decreases with age. Early assessment is therefore essential. Once stuttering becomes chronic, communication can be severely impaired, with devastating social, emotional, educational, and vocational effects.

The variability of stuttering behaviors is one of the facts about stuttering and something that contributes as much as anything else to the mystery of the disorder. It is difficult for listeners to understand how children who stutter (CWS) can be speaking fluently at one moment and a word or two later struggle dramatically as they attempt to do something as common as saying their name. The variability of stuttering behavior also makes it difficult for clinician to become accustomed to CWS, for it is not always possible to predict whether or not a person will stutter. The high degree of intra-speaker variability makes the assessment of fluency more formidable than it may first appear. The great variability of stuttering from time to time under different conditions is liable to result in assessments that are unrepresentative.

II. STUTTERING EVALUATION FOUNDATION FRAMEWORK

A. Problem Formulation

There are many therapeutic approaches that have been shown to help CWS. To be sure, the logic and techniques associated with most intervention strategies provide the clinician with a framework and a sense of direction about the syndrome and its treatment. Each strategy comes with its own doctrine. Each of these approaches can provide something of value for the clinician and her client, depending on such variables as the needs of the client, the stage of treatment, and the talent and experience of the clinician. Almost any therapy has the power to eliminate stuttering in someone, sometime, and someplace. The uniqueness of each individual clinician and each of the CWS prevents any specific recommendations of therapy techniques from being universally applicable [2].

Stuttering has varied treatment techniques, only a few have been tested for their efficacies. Some are questionable; some have uncontrolled clinical support; several are purely rational. Most clinicians combine certain components of techniques to create somewhat personal programs. The integrated or individualized approach combines fluency shaping and stuttering modification, based on the client's age and other key characteristics. Airflow management, gentle phonatory onset, and rate reduction through prolonged syllables are common elements across diverse contemporary treatment programs, these programs are empirically supported with experimental evidence [3].

The scores from manual measures do provide quantitative information that assists clinician in supporting clinical decisions and verifying progress. This quantitative measurement may be varied from one clinician to another which leads to the scoring inconsistencies. Scoring generated by computer-based system enables clinicians in different locations to be on "the same page" regarding the general severity and overall characteristics of clients. The utility of any measurement device will be influenced by such client characteristics as motivation, honesty, intellectual ability, and educational background. The choice of a scale is determined by the treatment strategy being used.

Treatment offered by clinician involves various techniques to slow down the speech, coordinate speech production with breathing, change the way of speaking and pronouncing words. It also involves counseling and relaxation therapy, which very often overlaps with work and input from the psychologist. These techniques work to certain

degree and the results depend very much on the particular case. Usually, they work better in the clinical setting than in the real world, where person cannot concentrate as much on speech production. Unfortunately, many of these techniques require a conscious effort on the part of the clients. In other words, the success of therapy usually depends a great deal on the amount of effort the client.

Not all CWS make as much progress as we would like, but almost any client who is ready and motivated enough to face his communication problem had-on is likely to make some progress during treatment. Therefore, element of motivation should be integrated into stuttering therapy. CWS often finds traditional therapy to be a very tedious and undesirable process. Successful intervention also requires continued commitment and motivation by CWS. The problem could be eliminated if clinicians use a "user-friendly" stuttering relief approach where this could be enhanced by the use of computer-based system. An experienced guide such as computer-based scoring analyses can show the way or, at the very least, make the journey more efficient and often more pleasant.

B. Implementation Design and Principles

Many of fluency-producing activities involve combinations of altered vocalization or enhancement of the speaking rhythm [4]. Fluency shaping is also referred as fluency modification. The essence of fluency shaping is the establishment of fluent speech in a controlled clinical setting which effectively and durably replaces the chronic stuttered speech pattern with a newly learned prolonged and rhythmic fluent speech. Once fluent speech is attained, it is shaped and expanded so that CWS can gradually maintain fluency in conversational speaking situations both within and outside the clinical setting.

The clinical program design follows and constitutes a major part of the present evidence-based practice (EBP) or treatment with CWS [5]. The concepts of establishment, transfer (generalization, out-of-clinic), and maintenance (over a long term period) must be employed. Follow-up post treatment data is collected to determine the positive, long-term effects of the programs. All these procedures have important contingency management features.

Fluency shaping procedures provide the CWS with techniques for both initiating and enhancing his fluency. The clinician cannot always assume that because the child's speech is nonstuttered, it is necessarily fluent. Speech that is to be expanded and reinforced should have high-quality fluency, which is characterized by smooth and effortless production. Regardless of the names for these therapy techniques, they consists of procedures to help the CWS more efficiently manage the breath stream, produce gradual and relaxed use of the vocal folds, use a slower rate of articulatory movement, make gradual and smooth transitions from one sound to another, produce light articulatory contacts, and keep an open vocal tract in order to counteract constrictions resulting from tension.

Our software based upon the physical analysis of speech sounds as they are being uttered. It provides real-time measures of sounds, evaluate the sounds against standards for their production, and immediately signal the results of the evaluation in graphs plotted on the computer screen. The implementation of computer-based evaluation system provides a faster way than traditional assessment process. Clinician can facilitate each client's progress by assessing client's history file. A log of the client's scoring is maintained in a personal file saved in computer. In addition, the personal score files maintain a count of the total number of times each utterance was practiced.

Three therapy techniques were introduced in computer-based method. The techniques are Shadowing, Metronome and Delayed Auditory Feedback (DAF). Clinical experience [6] showed that CWS are also likely to stutter less when they repeat words they hear. A task referred to as shadowing, in which a speaker repeats what he/she hears as quickly as possible, reduces stutter frequency significantly. Shadowing is the spontaneous speech equivalent of reading in chorus.

Metronome and automatic speech induce greater fluency in the CWS. They are encouraged to speak syllable by syllable, with each syllable stressed evenly, spoken in a regular rhythm and separated equidistantly from the next syllable [7]. Doing so will cause CWS to concentrate on how they are speaking and thus reduce their speaking rates.

DAF has been documented to improve fluency in CWS. The increased fluency has been attributed to the slowed speech rate induced by DAF. The DAF technique uses a tape-recording device that reproduces the voice through earphones after a few milliseconds. Most CWS stutter less severely while speaking under condition of DAF with a delay of 250-millisecond [8].

Rewards are important to motivate CWS. Research [9] supported that both tangible forms of rewards and verbal rewards were effective in reducing stuttering. Both forms of reward appeared to be successful, but their unique contributions could not be measured because treatment involved a number of therapy procedures. It is essential that the child enjoys the assessment process and finds it to be a positive experience. Therefore, we implement applause compliments and firework displays for CWS who managed to obtain a score of 80 and above.

III. METHODOLOGY

A. System Design

Sound recording of 5 goal utterances is implemented where the software incorporates functions record, playback, open, close and save of standard WAVE file. The 5 goal utterances are customized by the clinician for each client depending on the age and language level. Each goal utterance is stored in a separate WAVE file, which can be individually selected for practice. The duration for each target utterances is six seconds. Background noise level is identified for each client's environment. In the speech pathology clinic, after identifying the client's stuttering problem, the clinician verbally records 5 speech utterances for client to practice during the assessment process.

During the process, the client selects playback to listen to the clinician's pre-recorded utterances. The client speaks an utterance into a microphone. The client practices matching the clinician's speech pattern via both audio and visual

means. The client can audibly and repeatedly listen to the target utterance by selecting playback. The visual comparison is achieved via the display of average magnitude profiles (AMPs) of both the clinician and client utterances on the same axis. The software is able to calculate and display the client's AMP as it is spoken.

AMPs are displayed for client to copy as closely as possible and it conveys to the client those locations where the client's spoken utterance differed from the clinician's signal in the aspect of amplitude, duration, onset, and end location. The speech processing is done in real-time. Real time display of AMP is very important as it allows the client to instantly evaluate and compare their speech to that of the clinician. The client can then, if necessary, alter their speech as required to closely match the clinician's AMP. This gives immediate feedback to the client's performance relative to the goal, and allows the client to anticipate what amplitude or rate change that is needed to reach the goal.

The scoring algorithms evaluate the client's performance where the scoring routines compare the client's utterance to the reference utterance in four categories. The categories are start location identification, end location identification, maximum magnitude comparison and duration comparison. Upon completion of each practice, scores are assigned to each trial. The scores are displayed to the client, allowing the client to observe the progress being made.

Software generates a history file summarizing the client attempts. A separate history file is created for each client. A log of the client's scores is maintained in the personal file on the client's computer. In addition, the personal score files maintain a count of the total number of times each utterance was practiced. Clinician can use this information to determine suitable therapy techniques for each client. Moreover, it enables clinician to assess or monitor client progress and observe how much time client spends to practice. Based on client's progress record, the supervising clinician is well-informed on which displayed signals and measured parameters would be useful for improving the speech rehabilitation process. These features enable the client to be success in future therapy sessions.

B. Design Detail

DC offset is the average vertical offset from 0dB that is in the recorded wave form. DC offset must be removed in order for the particular AMP to appear the same on all computers. The same speech waveform containing different DC offsets will result in two unique AMPs if the DC offset is included in the calculation of the average magnitude. The DC offset is calculated by determining the average value for each 25 ms segment of speech. DC offset is then subtracted from each sample value as shown in (1), which allows the level shifting of the speech signal back to zero.

$$\text{Average Magnitude Profile} = \frac{\sum_{t=0}^{t=25ms} |\text{Speech Sample} - \text{DC Offset}|}{400} \qquad (1)$$

WAVE files are processed using analysis and synthesis window lengths of 400 samples with Blackman windowing. Blackman is chosen because it gives the best attenuation

which means that it is clear to hear an audible difference between filtered and non-filtered samples [10]. The Blackman window offers a weighting function similar to the Hanning but narrower in shape. It has all the dynamic range any application should ever need. Blackman window equation is shown in (2).

$$w(n) = 0.42 - 0.5\cos\left(2\pi\frac{n}{N}\right) + 0.08\cos\left(4\pi\frac{n}{N}\right) \qquad 0 \le n \le N \qquad (2)$$

The window $w(n)$ is chosen to have a duration of 25 ms or 400 samples. The 25 ms window is shifted by 10 ms steps. Therefore, a new window of 400 samples is calculated every 10 ms. A 25 ms window $w(n)$ duration is a common window duration. It is sufficient to capture all the stuttering disfluencies. The four disfluencies of stuttering are syllable repetitions, word repetitions, prolongation of a sound and blocking or hesitation before word completion. Calculation of AMP requires the calculation of 600 average magnitude values corresponding to the 600 frames in 6 seconds of speech data. The average magnitude is calculated for each 25 ms of speech data with a new average magnitude calculated every 10 ms.

Time-domain filtering is favored because the evaluation system is a real-time application in which it is important to process a continuous data stream and to output filtered values at the same rate as raw data is received [11]. Time domain filters are used because the desired information is encoded in the shape of the signal's waveform. Time domain filtering performs the function of low pass filter (LPF) and reduces the sampling rate from 16 kHz to 100Hz. The filter thus removes all frequencies above the Nyquist frequency and prevents aliasing. Samples of 401 and onwards are filtered out to eliminate the undesired high frequency components.

Modern operating systems such as Windows provide a quite useful Application Programming Interface (API) for programming soundcards. The normal way of outputting audio is to open a device and writing blocks of data to this device. The audio data is written to output buffer. The output buffer is a block of memory which has several constrictions. The data in this buffer is usually transferred to the soundcard using the Direct Memory Access (DMA) controller. The DMA controller is a device which can copy data between memory and hardware devices without needing the CPU. Sound input works generally in the same way as the output except in opposite direction.

IV. RESULT ANALYSIS

Test subjects were selected from 6 primary schools located in Skudai, Malaysia. A total of 10 subjects participated, 10 males and 1 female. The age span was between 8 and 12 years old.

The speech sample was presented to the clinician in both the audiovisual and audio-only mode. Clinician judged each sample individually. Table 1 shows the scoring generated by software for each subject. During clinical session, clinician assessed the subject by making the measurement of percentage of stuttered syllables (%SS). Software scoring was compared to clinician's %SS as shown in Figure 1. The data indicated that clinician has agreed on all the therapy techniques determined by the scoring generated by the

software. The accuracy of the software is identified as 100%. This software has demonstrated great potential to aid clinician in determining suitable therapy technique for each of the CWS.

TABLE I
Software Scoring for Test Subjects

Subject	Scoring (%)		
	Technique Shadowing	Technique Metronome	Technique DAF
A	72	60	80
B	76	74	73
C	85	84	76
D	68	75	82
E	81	88	75
F	68	82	74
G	86	85	90
H	82	80	76
I	62	68	75
J	60	57	38

System effectiveness is the extent to which a system or software employed in the field does what it is intended to do for a specific population [12]. Based on the software and clinician analysis, the following are the common observation:

- Certain sounds are more likely to be stuttered than other sounds, mainly consonants, but with wide individual variations as to what particular sounds are problematic, although word-initial sounds are often a major determinant.
- Certain parts of speech are more likely to be stuttered than other parts of speech, videlicet, adjectives, nouns, adverbs and verbs such as words belonging to open word classes.
- The position of a word in a sentence affects the degree of difficulty it presents to the client, the first three words of a sentence being stuttered more often than words occurring later.
- Longer words seem to be stuttered more often than shorter words.

Subject	Software			Clinician		
	Technique Shadowing	Technique Metronome	Technique DAF	Technique Shadowing	Technique Metronome	Technique DAF
A			■			■
B	■			■		
C	■			■		
D			■			■
E		■			■	
F		■			■	
G			■			■
H	■			■		
I			■			■
J	■			■		

Fig. 1. Comparison between the Determination of Therapy Technique for Each Test Subject by Software and Clinician

V. CONCLUSION

Initiation of a computer-based stuttering evaluation system presents an important challenge and is one that the field must embrace in order to advance. It is critical that standards are adopted for evaluating therapy techniques so that significant variations in outcomes can be reliably identified. On the whole, future studies should focus on identifying specific therapy techniques that contribute the most to successful treatment outcomes as well as variables that are responsible for treatment failures. Our current software aimed at determining suitable therapy techniques for each client by looking at the scoring generated automatically by the software. This personal computer-based system is intuitive to use, cost-effective, and easily integrated into current speech rehabilitation regimen. We believe that our software tool will improve the effectiveness and availability of stuttering treatment.

REFERENCES

[1] Melissa Peter. STUT (E): Stuttering. *Malaysian Association of Speech Language and Hearing*. 2002.
[2] Harold B. Starbuck, *Therapy for stutterers*, Memphis, Tennessee: Stuttering Foundation of America, 1992, pp. 45 – 51.
[3] Patrick Finn, "Evidence-based treatment of stuttering: ii. clinical significance of behavioral stuttering treatments," *Journal of Fluency Disorders*, vol. 28, pp. 209-218, 2003.
[4] Katrin Neumann, Christine Preibisch, Harald A. Euler, Alexander Wolff Von Gudenberg, Heinrich Lanfermann, Volker Gall and Anne-Lise Giraud, "Cortical plasticity associated with stuttering therapy," *Journal of Fluency Disorders*, vol. 30, pp. 23-39, 2005.
[5] Sackett, D. L. and J. Sonis, *Evidence-based medicine: How to practice and teach EBM*, Edinburgh, New York: Churchill Livingstone, 2000, pp. 250.
[6] Kalinowski, J., Saltuklaroglu, T., Guntupalli, V, & Stuart, A., "Gestural recovery and the role of forward and reversed syllabic repetitions as stuttering inhibitors in adults," *Neuroscience Letter,* vol. 363, pp. 144-149, 2004.
[7] Robert Eklund, *Disfluency in Sweedish human-human and human-machine travel booking dialogues*, Ph. D. Thesis, Linkopings University, 2004.
[8] Radford, N., Tanguma, J., Gonzalez, M., Nericcio, M.A., Newman, D., "A case study of mediated learning, delayed auditory feedback, and motor repatterning to reduce stuttering," *Perceptual and Motor Skills*, vol. 101, pp. 63-71, 2005.
[9] Thomas, C., & Howell, P., "Assessing efficacy of stuttering treatments," *Journal of Fluency Disorders,* vol. 26, pp. 1-23, 2001.
[10] Don Morgan. *Practical DSP modeling, techniques, and programming in C*. Canada: John Wiley & Sons, Inc. 1997.
[11] William H. Press, Brian P. Flannery, Saul A. Teukolsky, William T. Vetterling, *Numerical recipes in C: The art of scientific computing,* 2nd ed., UK: Cambridge University Press, 1992.
[12] Last, J. M. A Dictionary of Epidemiology. New York: Oxford University Press. 1983.